Universiteit Utrecht

Faculteit Geesteswetenschappen

UiL OTS

MA Thesis

Research Master Linguistics

Force Dimensions in Dutch Manner Adverbs:

A self-paced reading study into inferences

Suzanne Dekker

3489337

Utrecht, 2016

Supervisors:

A. Goldschmidt, MA

Dr. J. Zwarts

CONTENTS

ABSTRACT

The goal of this thesis is to find empirical evidence for the subdivision of manner adverbs into mono-dimensional adverbs such as *lightly*, directly modifying a quality dimension of the verb; and adverbs such as *friendly*, that target a more complex dimension. Importantly, complex adverbs like *friendly* can also modify basic quality dimensions via a cancellable inference. Evidence for this type of inference is presented in the form of questionnaires and a self-paced reading study. These inferences can be accounted for in Neo-Davidsonian Event Semantics by means of additional conceptual information in the form of meaning axioms.

Key-words: adverbial modification, defeasible inferences, self-paced reading experiment, Neo-Davidsonian Semantics, conceptual knowledge

1. INTRODUCTION

One of the means of modifying a verb is to make use of an adverb. This class of words is highly flexible with respect to their distribution and inferential properties (cf. Maienborn & Schäfer 2011). Recent literature on adverbs suggests several subclasses of this category on the basis of properties of the verb or sentence they specify, as well as their position in the sentence (e.g. Hofland 2011, Maienborn & Schäfer 2011, Schäfer, 2013). According to their grammatical function, adverbs belong to the wider group of adverbials, which typically refer to the time, place or manner of an event, as in (1).

1. a. Paul laughed *the whole day*.

   b. The children played *in the garden*.

   c. Henriette dances *beautifully.*

(Maienborn & Schäfer, 2011: 2)

Maienborn and Schäfer show that the example in italics in (1a) is a noun/determiner phrase, (1b) a prepositional phrase and (1c) an adverb phrase, yet the phrases all share the function of adverbials. Furthermore, (1a) modifies the temporal dimension of an event and specifies *when* an event took place; (1b) modifies the location and specifies *where* an event took place; (1c) modifies the manner, and specifies *how* the event took place.

The latter category, manner adverbs, is the focus of our current study. This class consists of several adverbs which appear to have in common that they answer the question *how* an event took place, but do so in several different ways. The following sentences in (2) contain adverbs in *italics* that are typically considered manner adverbs:

2. a. Peter runs *fast/slowly.*

   b. Marie sings *loudly/quietly*.

   c. Kim dances *beautifully/woodenly*.

(Maienborn & Schäfer, 2011: 24)

Although the examples in (2) are all manner adverbs, the meaning they add to the sentence differs. Maienborn and Schäfer show that the standard paraphrases for manner adverbials, *the way X VERBs is ADJ*, is not appropriate for every item in (2). While it is possible to paraphrase (2c) in this way, i.e. to say that *the way Kim dances is beautiful*, for (2a) or (2b) this is an awkward paraphrase. In (2a), *to run fast* means that the speed of the event was fast, and not that the 'manner' in which Peter runs. Similarly, in (2b) *loudly* modifies the volume of the singing and not the 'manner' (cf. Maienborn & Schäfer, 2011: 24-25). There are in these cases several *dimensions* to an event that can be modified by manner adverbs, such as speed and sound.

While the adverbs in (2a,b) modify primary dimensions of an event, e.g. speed or sound, the adverb in (2c) seems to modify a more complex dimension. We can therefore speak of mono-dimensional manner adverbs such as *fast* and *loudly*, and more complex or multidimensional manner adverbs such as *beautifully* (cf. Maienborn & Schäfer, 2011: 14-15).

Please note that the term manner is now used in two senses. Firstly, what we will call complex manner adverbs characterises a class of adverbs that elucidate how an event took place. Secondly, 'manner' (henceforth between single parentheses) is used as a category that is the target of modification for the more complex adverbs that modify a quality dimension which is not as clearly defined as e.g. speed or sound.

Modification by what we call complex manner adverbs is further explored in Geuder (2000), who shows differences in adverbs in (3) below.

3. a. Die Perlen glänzten matt / rötlich / feucht.
   The pearls were gleaming + modifiers: dull / reddish / moist.
   b. Das Licht war ?? feucht hell / ?? rötlich hell.
   The light was moistly / reddishly bright.
   c. Die Lampe leuchtete hell / rötlich / ?? feucht.
   The lamp was shining brightly / reddishly / moistly.

(Geuder, 2000: 117)

While *glänzen* (to gleam, glow) in (3a) can be paired with the modifiers *matt* (dull), *röttlich* (reddish), and *feucht* (moist), the verb in (3c) can only be modified by some adverbs such as *hell* (brightly) but not others. In (3b), it can be seen that the adjective *hell* (bright) does not allow for modification by the adverbial predicates in (3a). Geuder uses these examples to show that adverbs can differ with respect to which predicates they can modify (cf. Geuder, 2000). In other words, an event has to be specified for a certain quality dimension, such as luminosity (brightness), colour (reddish) and humidity (moist), in order to be modified by adverbs that target those different quality dimensions.

What we are specifically interested in is manner adverbs that modify the force dimension of an event. For instance, the adverb *lightly* in sentence (4b) lessens the amount of force used in the hitting compared to (4a), while (4c) increases the amount of force used.

4. a. Nancy hit Oliver.
   b. Nancy hit Oliver lightly.
   c. Nancy hit Oliver hard.

For sentences (4b) and (4c), the paraphrase *Nancy hit Oliver in a light/hard manner* is equally awkward as in the examples in (2). In (4a,b), *to hit lightly/hard* means that the force of the hitting was decreased/increased with respect to the prototypical hitting event, respectively. From the paraphrase, we conclude that these manner adverbs are mono-dimensional, and that they modify the force dimension of an event. Conversely, Goldschmidt (2013) illustrates examples of adverbs that modify the force dimension of an adverb only by means of an inference in sentence (5) below (2013: 10):

5. a. Nancy hit him playfully but still rather hard.
   b. # Nancy hit him playfully but still rather lightly.
   c. # Nancy hit him lightly but still rather hard.

In sentence (5) we make use of Lakoff's denial-of-expectation (Lakoff, 1971). In short, the conjunct *but* encodes a denial-of-expectation meaning between the two conjuncts in the sentences. The expectation that is denied in sentence (5a) is therefore the opposite of the force-increase adverb *hard,* which means that the force-decreasing *lightly* is expected. On the other hand, *playfully* in (5b) results in an infelicitous sentence when in contrastive conjunction with an adverb that decreases force. We therefore consider *playfully* to have a force decrease reading in the context of *to hit*.

Furthermore, where *playfully* in (5a) can be in a contrastive conjunction with *hard*, which directly increases the force of the hitting event, *lightly* in (5c) cannot. This shows that

playfully does not express a decrease in force via its lexical semantics, but only indirectly modifies the force dimension by means of an inference (Goldschmidt 2013: 4). The adverb *playfully* therefore modifies at least two dimensions of an event: it directly modifies the complex dimension of manner, and indirectly also modifies the force dimension by means of an inference.

Similarly, this observation can be made with adverbs in increasing force direction such as *angrily*. Sentence (6) illustrates how *angrily* in the context of *to hit* results in a cancellable inference on the force dimension, similar to *playfully*.

6. a. Nancy hit Oliver angrily.

   b. ✓ Nancy hit Oliver angrily, but still rather lightly.

   c. # Nancy hit Oliver angrily, but still rather hard.

A distinction is therefore created between adverbs such as *lightly* and *hard*, which directly modify the force of an event and are therefore considered mono-dimensional; and *playfully* and *angrily*, which can modify the force of an event indirectly, via a cancellable inference. We therefore call the latter category multi-dimensional adverbs. These multi-dimensional adverbs can modify several dimensions of meaning of an event: the complex manner dimension by virtue of their lexical semantics, and the force dimensions via an inference. This leaves us with the question of how to analyse the multi-dimensional adverbs. This thesis will attempt to answer that question by elaborating on multi-dimensional adverbs and exploring them by means of a self-paced reading study.

The aim of this thesis is to discuss the domain of adverbial modification, and more specifically multi-dimensional inferences within adverbial modification. This will be discussed in the Neo-Davidsonian paradigm as this framework offers a straightforward way of accounting for adverbial modification, namely analysing adverbial modifiers as simple first order predicates that add information about an event (cf. Maienborn, 2003). The way this will be done is to focus on a specific domain of application, i.e. the force dimension, modified through inferences triggered by multi-attribute adverbs such as *playfully* and *angrily*. The structure of this thesis

will be as follows: in chapter two I will discuss the nature of this phenomenon more elaborately; chapter three will provide empirical evidence as to whether the inference in (6) is truly a separate inference step by means of a set of questionnaires and a self-paced reading experiment; chapter four will discuss the results of this experiment and its theoretical implications in a Neo-Davidsonian framework. Finally, chapter five will summarize, discuss and conclude this thesis.

2. ON THE NATURE OF *FRIENDLY* INFERENCES

Although the phenomenon described in the previous section has loosely been called an inference, it is not entirely clear what we are dealing with. The question arises whether we observe an entailment, a presupposition, an implicature of some type, or some other defeasible inference. This section aims to clarify which of these we are dealing with.

It can easily be seen that our phenomenon is not an entailment, as entailments are not defeasible. Sentence (7) shows via the denial-of-expectation test (cf. Lakoff, 1971) explained in the previous chapter that *friendly* modifies the force dimension, but this force modification is subsequently cancelled by the predicate *hard* in a contrastive conjunction. Our phenomenon is clearly defeasible, and therefore not an entailment.

7. Nancy hit Oliver friendly but still rather hard.

We can rule out presupposition equally fast by testing for one of the features of presuppositions, namely survival of the presupposed meaning under projection, e.g. from a conditional clause. This is exemplified in the sentences (8) and (9) below:

8. a. Nancy hit Oliver friendly.
   b. Nancy hit Oliver lightly.
9. a. If Nancy hit Oliver friendly, he won't be upset.
   b. Nancy hit Oliver lightly.

Sentence (8a) contains an adverb considered to be multi-dimensional: *friendly* gives rise to an inference in the force dimension when combined with a hitting event, as seen in (7). This inference is given in sentence (8b). In sentence (9a), the phrase that should trigger the inference is contained in an if-clause. The expected inference that should arrive is given in (9b). However,

it does not survive projection from the if-clause, and the inference does not arise in (9a). We can therefore conclude we are not dealing with a presupposition.

Our phenomenon is also not a conventional implicature, because those are not cancellable (cf. Horn, 2004). What remains as candidates are generalized conversational implicatures, and defeasible inferences. The line between the two is complex: Gricean generalized conversational implicatures (hence: GCI's) arise due to a speaker violating, exploiting or flouting a conversational maxim. Inferences, on the other hand arise by a cognitive process "by which participants figure out (speaker) meaning beyond what is said or encoded" (Haugh 2013: 2). In other words, there are two ways of looking at the same thing. Implicatures are considered the speaker's intended meaning; if the hearer recognises this intent, it becomes an inference.

GCI's are presumed meanings that are the result of rational, communicative behaviour, and make use of Q, I and M heuristics to arrive at the presumed meaning (Jaszczolt 2014; Levinson 1995, 2000):

Q: What isn't said, isn't.

I: What is expressed simply is stereotypically exemplified.

M: What's said in an abnormal way isn't normal

These heuristics are an extension of Horn's (1984) Q and R principles, with Levinson's and Jaszczolt's M principle as an interaction between Horn's Q and R (cf. Horn, 1984). The heuristics above can be illustrated by the following examples, taken from Carston (2004: 181):

10. a. Some of the children passed the test

    b. Some **but not all** of the children passed the test.

11. a. Mary looked at John and he smiled.

    b. Mary looked at John and **then** he **=John** smiled.

12. a. Nick was instrumental in lighting the fire.

    b. There was **something odd in the way** Nick lit the fire.

Sentence (10) shows the classic example of the scalar Q-heuristic in relation to a scale of lexical alternatives, e.g. <some, all>. Assuming that the speaker would give the most informative sentence, and did not say that all children passed the test, it must mean that the alternative was only some children passed the test. The I-heuristic is at work in sentence (11), as conjoined sentences stereotypically give information in a sequential order . The M-heuristic is shown in sentence (12), where a deviant expression with respect to the unmarked sentence *Nick lit the fire* is used to indicate that there was a way in which Nick acted abnormally when lighting the fire.

Similarly, the I-heuristic seems to explain the phenomenon in (13):

13. a. Joan hit Harry friendly on his shoulder.

    b. Joan hit Harry friendly **=lightly** on his shoulder.

    c. Joan hit Harry friendly **=lightly**, but still rather hard **=not lightly**, on his shoulder.

In (13a), the added meaning of *friendly = lightly is* due to the assumption that a stereotypical friendly event would encompass friendliness and therefore a low force, as friendly events stereotypically involve light force. Nevertheless, this effect can be negated, as seen in sentence (13c). However, Levinson (1995, 2000) claims that this is the correct approach for presumptive meanings that arise on utterance-level, but not those that can arise on word-level (cf. Levinson 1995, 2000; Jaszczolt 2014). Levinson (2000) agrees that the correct way of arriving at a standard Gricean GCI for an utterance token is via the Q, I and M heuristics; however, he argues that the pragmatics of the presumptive meaning is a matter of preferred or default (presumptive) interpretations. These presumptive interpretations are added via the structure of utterances instead of by the particular contexts of utterances. Importantly, they are based not on the speaker's intention, but on general expectations on how language is normally used (Levinson 2000: 22) and the hearer need not have processed everything before "jumping to

conclusions" (cf. Jaszczolt 2014). In other words, presumptive interpretations are automatically generated based on default usage of certain expressions and structures.

Carston (2004) concludes that the conjunction "and" in (11a) "carries a default rule to the effect that the event described in the first conjunct preceded that described in the second conjunct" (2004: 182). Since these inferences are by default, they can be overridden if their context is inconsistent with the default output. With regards to our multi-attribute manner adverbs, one could conclude when hearing that a hitting event is done in a friendly manner, the hearer expects that the event in combination with *friendly* results in a lessened amount of force, even though this may subsequently be cancelled by context.

Overall, the defeasibility of the phenomenon in (13c), especially in context of the I implicatures explained above, leads me to classify it as being a form of GCI, which is an inference. Throughout the literature, the terms presumptive meaning, GCI and inference are used interchangeably (cf. Levinson 1995, 2000; Jaszczolt 2014, Carston 2004). Henceforth, I will therefore for simplicity's sake describe the defeasible effect of friendly hitting events observed in the first section as an inference.

Finally, a problem remains, namely an issue with the cancellation of presumptive meanings. As noted by Jaszczolt (2014), "it is difficult at present to decide between the rival views that (i) a particular GCI arose and was subsequently cancelled or (ii) did not arise at all due to being blocked by context" (2014:4). In order to differentiate between the two possibilities, we set up an experiment to show whether the inference is triggered and subsequently cancelled during on-line processing. This will be discussed in the next chapter.

3. SELF-PACED READING EXPERIMENT

As shown previously, there are adverbs that can be considered multi-attribute such as *playfully,* which indirectly have an effect on the force dimension by means of an inference. This study aims at testing native Dutch adults to see whether the inference is indeed drawn and can subsequently be cancelled during on-line processing. It investigates the possibility of an increased cognitive workload due to the process of cancelling an inference. This will be measured in a self-paced reading task: participants decide the speed with which they read sentences on a computer screen by pressing a button and making the words appear one-by-one. The expectation is that participants will have an increased reading time in the stimuli sentences due to the added cognitive load of having to cancel an inference.

3.1 METHODS

In order to create the required stimuli to prove our inference exists, we first needed a list of single-attribute and multi-attribute adverbs in combinations with several verbs of contact. The first questionnaire was launched to find suitable stimuli sentences.

First of all, we tested whether verbs of contact allow for both force directions by combining them with benchmark adverbs:

14. a. Jet sloeg Harry **hard** op zijn schouder. (force increase)
    Jet hit Harry hard on his shoulder.
    b. Jet sloeg Harry **zachtjes** op zijn schouder. (force decrease)
    Jet hit Harry lightly on his shoulder.

We translated the German force denoting adverbs from Goldschmidt (2013), and combined the translations with our benchmark verb of contact *slaan* (to hit) in contrastive conjunction with our benchmark adverbs.

15. a. #Nancy sloeg Oliver zacht maar toch vrij hard.
    Nancy hit Oliver softly but still rather hard.
    b. Nancy sloeg Oliver speels maar toch vrij hard.
    Nancy hit Oliver playfully but still rather hard.
    c. # Nancy sloeg Oliver speels maar toch vrij licht.
    Nancy hit Oliver playfully but still rather lightly.

The contrastive conjunction was used for several reasons. Firstly, two single-attribute adverbs of opposite force directions cannot be in a contrastive conjunction, as in (15a) where # marks infelicity; secondly, multi-attribute adverbs can occur in this conjunction with single-attribute adverbs of an opposite force direction, as in the felicitous (15b), because there is a contrast; thirdly, two adverbs in the same force direction cannot be contrasted to each other, as in (15c).

The sentences in (15) make use of the denial-of-expectation test mentioned in the introduction. This allowed us to see whether an adverb could felicitously pair with our benchmark adverbs for force increase and decrease, and resulted in our classification of mono-dimensional and multi-dimensional adverbs. If an item can pair with our benchmark force increasing adverb, “but still rather hard”, the adverb was classified as “multi-dimensional decrease”; if an item can pair with our benchmark force decreasing adverb, “but still rather lightly”, the item was classified as “multi-dimensional increase”.

Several sentences of the format in (16b) were piloted in a questionnaire among 30 native Dutch speakers, who had to rate the acceptability of the sentences on a four-point Likert scale. The points ranged from 1 “clearly bad”, 2 “maybe bad”, “3 maybe good” and 4 “clearly good”. From this, we gathered possible stimuli in the sentence format in (16a). In order to proceed with

a self-paced reading experiment, we needed to compare the reading times of our stimuli to the reading times of control sentences in which there is no inference cancellation, as in (16b).

16. a. Nancy sloeg Oliver vriendelijk, maar toch vrij hard, op zijn schouder.
    Nancy hit Oliver friendly, but still rather hard, on his shoulder.
    b. Nancy sloeg Oliver licht, maar toch vrij hard, op zijn schouder.
    Nancy hit Oliver lightly, but still rather hard, on his shoulder.

Unfortunately, these control sentences were problematic: they are contradictory. Currently, we have no empirical evidence to suggest that processing of a contradiction would be faster, slower or simultaneous with the cancelling of an inference. We therefore changed the format of our stimuli and control sentences to that in (17).

17. a. Nancy sloeg Oliver vriendelijk, en niet zachtjes, op zijn schouder.
    Nancy hit Oliver friendly, and not softly, on his shoulder.
    b. Nancy sloeg Oliver licht, en niet zachtjes, op zijn schouder.
    Nancy hit Oliver lightly, and not softly, on his shoulder.

A second questionnaire was launched to test several items in the format in (17). In this format, the inference in (17a) is triggered on the first adverb, and cancelled by the negated second adverb in a coordinating conjunction. Again, our control sentences turned out to be unsuitable: the problem with this format was that the control sentences were tautological, and received poor judgment scores. These items were therefore not usable for a self-paced reading experiment. A third questionnaire was launched to test the items in the format in (18).

18. a. Nancy sloeg Oliver vriendelijk, en tegelijk vrij hard, op zijn schouder.
    Nancy hit Oliver friendly, and simultaneously rather hard, on his shoulder.
    b. Nancy sloeg Oliver boos, en tegelijk vrij hard, op zijn schouder.
    Nancy hit Oliver angrily, and simultaneously rather hard, on his shoulder.

In the sentences in (18), the coordinating conjunction was used; there no longer is a denial-of-expectation in the conjunction "and" instead of "but". Instead, the denial-of-expectation is introduced via *vrij* ("rather"), which leaves the sentence still suitable for cancelling our inference. In order to avoid a temporal interpretation wherein Nancy first hit Oliver friendly, and subsequently rather hard, we introduced simultaneity through "en tegelijk" (and simultaneously). In this format, the inference is still triggered in (18a), and subsequently cancelled. For the control sentences, there is also an inference triggered on the first adverb; this inference, however, remains throughout the sentence as the second adverb in (18b) does not cancel it. As a result, the control sentences in this format were mildly tautological, but considered acceptable by participants of the third questionnaire.

The three questionnaires show that there is a force inference arising on adverbs such as *vriendelijk* (friendly), and that this is indeed an inference and not part of the lexical semantics of the adverb. Having found a suitable format for our experiment, we continue with our self-paced reading task to determine whether the cancellation of the inference occurs during on-line processing.

3.1.1 MATERIALS

The materials that have been used for this study have been designed along the lines of the stimuli and controls for the third questionnaire. All items were previously tested for acceptability. In total, there were be 12 stimuli sentences with multi-attribute decrease adverbs combined with a single-attribute increase adverb, as shown in (19a); and 18 stimuli sentences with multi-attribute increase adverbs combined with a single-attribute decrease adverb, as shown in (19c). The corresponding control items made use of a multi-attribute adverb in the first position in order to compute the inference, combined with single-attribute items in the same force direction so that the inference is not cancelled. This can be seen in (19b-d). Again, the

last four words of the sentences are exactly the same so that any reading time variation would not be affected by those words but only by the adverbs.

19. a. Jet sloeg Harry speels, en tegelijk vrij hard, op zijn schouder.
    Jet hit Harry playfully, and simultaneously rather hard, on his shoulder.
    b. Jet sloeg Harry boos, en tegelijk vrij hard, op zijn schouder.
    Jet hit Harry angrily, and simultaneously rather hard, on his shoulder.
    c. Jet sloeg Harry boos, en tegelijk vrij zacht, op zijn schouder.
    Jet hit Harry angrily, and simultaneously rather lightly, on his shoulder.
    d. Jet sloeg Harry speels, en tegelijk vrij zacht, op zijn schouder.
    Jet hit Harry playfully, and simultaneously rather lightly, on his shoulder.

The verb in the sentence was also expected to influence the judgments. Out of 15 verbs, 12 require a minimal force, for example *slaan* (to hit), which is coded into the lexicon; 3 verbs were considered neutral with respect to force, for example *plaatsen* (to place). All verbs were tested for possible force increase and decrease inferences before the first questionnaire. Although this could mean that force increase inferences are harder to cancel with the hit-type verb because of the verbs' prototypical force. Due to this, the ratings for inferred force increase sentences were expected to be slightly lower (2-3) than the ratings for inferred force decrease sentences (3-4).

All sentences will be rotated in a Latin Square grid: no participant will see corresponding control items and stimuli. There were twice as many filler sentences as stimuli and controls combined in order to prevent the participant from guessing what the experiment is about and formulating a clear strategy that prevents intuitive answering.

3.1.2 PARTICIPANTS, PROCEDURE & DESIGN

The reading time experiment was conducted with 91 native speakers of Dutch recruited from the general vicinity of Utrecht. To make use of a Latin Square design, participants were divided over two lists: 45 participants in list 1, and 46 participants in list 2. No participants with SLI, dyslexia or other cognitive developmental difficulties were allowed to complete the test. Participants received a compensation of €5,- for their effort.

The experiments were conducted in a sound-proof cabin, and made use of a computer screen, a two-button response pad and a mouse. Participants were instructed to first read the sentences word by word, progressing by means of a button press. They were asked to use only one of the response pad buttons for this. After each sentence, the sentences had to be on a four-point Likert scale ranging "clearly bad", "maybe bad", "maybe good" and "clearly good". There was a 25% chance for a yes/no comprehension question following each sentence to encourage careful reading.

Every participant first was presented with a 7-sentence trial round to familiarize them with the manner of reading and the judgment task. In the trial round, the chance of a comprehension question was raised to 100%. After each trial round, the experimenter would enter the cabin to see if any further clarification was needed.

The experiment was programmed using ZEP. Items were presented in a pseudo-randomized order so that no more than two items of the same type (increase, control increase, decrease, control decrease, contradiction, correct, obviously correct) appeared consecutively.

3.2 RESULTS

Both the reading time and sentence judgments were recorded, and will be discussed separately. First, I will give an overview of how we prepared our data for analysis. Our main interest lay in the reading time analysis; this will be discussed first in section 3.2.2. Subsequently, the rating analysis will be briefly discussed in section 3.2.3.

3.2.1 DATA TREATMENT

Before analysing the data, two items were removed from the dataset: one filler and one stimulus item. The filler item was one of the contradictions, and was consistently rated on the high end of the scale. It was therefore considered questionably contradictory; and the item was excluded. The stimulus item removed was the particle verb "aandrukken", in English "to press (down)". The reason for this was an oversight in the experimental design: this verb would have its particle "aan" in the spillover area of the critical adverb. Because of this, once the particle verb appeared there could be a re-computation of the verb structure triggered. The oversight was that this also resulted in four words in the spillover area instead of three, which affected the average reading times and made it impossible to compare to the other items, which only had three words in the spillover area. The item was therefore discarded.

Several participants were also removed from the dataset because they showed a high error rate in the control questions. Our standards for the accepted level of correct responses was set relatively low for such a task, namely 70%, because of an intervening judgment task elongating the time between reading the sentence and answering the yes/no question. There were 9 participants in total, divided over both lists, that scored less than 70% of their control questions correct; they were therefore excluded.

On top of the control questions, we also applied scoring mechanism to judge the acceptance rate of the contradictory filler sentences. Rating a contradiction with a 'clearly bad' rating, the lowest on the 4-point scale, would award 1 point to the participant; the highest rating, 'clearly good', would award 4 points. As there were 19 contradictions remaining after the exclusion of one, the minimum score would be 19; a lenient score over 30 points resulted in the exclusion of another 9 participants. This was done to remove any participant that was overthinking the sentences until they seemed correct, and whose intuitions would no longer be trusted due to a yes-bias.

The effects of inference cancellation on reading time might not be immediately visible on the critical adverb, but carry over to the words immediately following it. As such, the reaction times of interest to us were the critical adverb and the three subsequent words in the spill-over area. They were named RT1, RT2, RT3 and RT4. For every RT, observations faster than 100 milliseconds were excluded, as this is the minimum amount of time needed between reading an item and activating the motor control function required to press the button for the next word (cf. Wehlan, 2008). Outliers were removed above 5000ms for RT1, 2000ms for RT2 and RT3, and 6500ms for RT4. For the final reaction time, RT 4, we accepted such a high latency because it was in sentence-final position, and participants were aware of the following judgment task which could make them hesitate. Finally, a log transformation was also applied to the data.

3.2.2 READING TIME ANALYSIS

In the following graph, the mean reading time data of the critical adverb and the three words following it is summarized. The four reading time means are RT1 for the critical adverb, RT2 for the preposition, RT3 for the article and RT4 for the final noun.

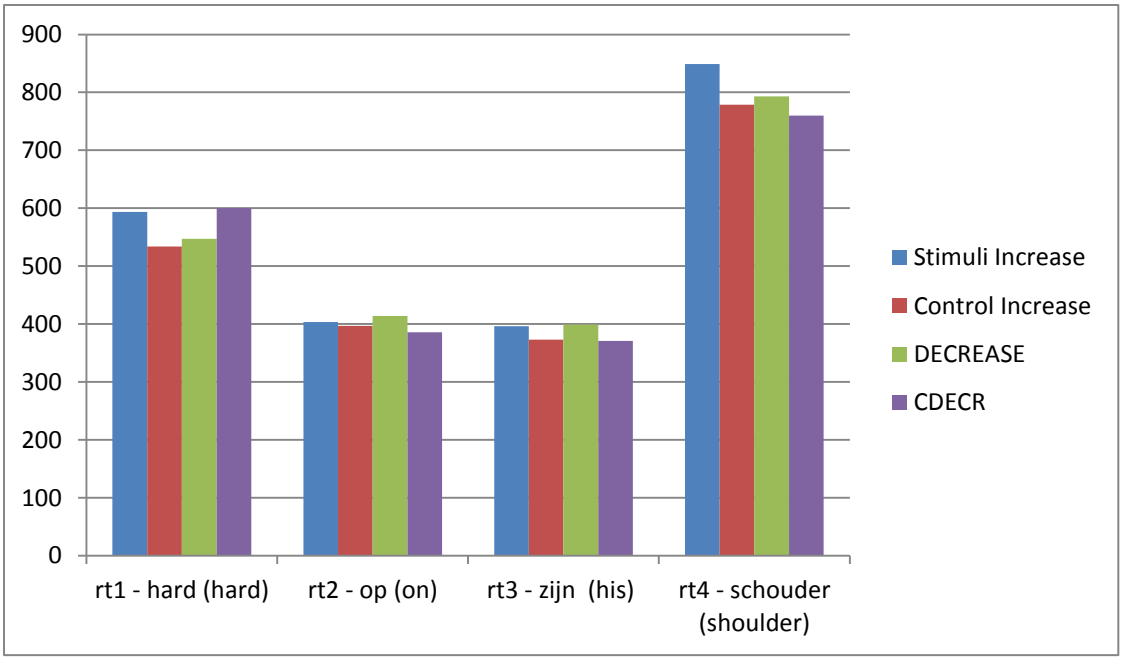


figure 1: mean reading times (all conditions)

We can see in figure (1) that the mean reading time for the critical adverb (RT1) was slower for the stimuli items in the increase condition, but not in the decrease condition. For the spillover areas RT2, RT3 and RT4, all stimuli items in both increase and decrease conditions were read slower than the control items. What is also interesting to note is that the average reading time for all conditions increases highly in RT4, with the increase stimuli condition highest of them all. This increased reading time could be due to the word in RT4 being a new noun instead of the same prepositional phrase as in all other stimuli. What could also be affecting these responses is that the sentence ended after RT4 and was immediately followed by the judgment task; there could be a wrap-up effect of the sentence, as well as hesitation due to participants being uncertain which answer they will give on the rating task.

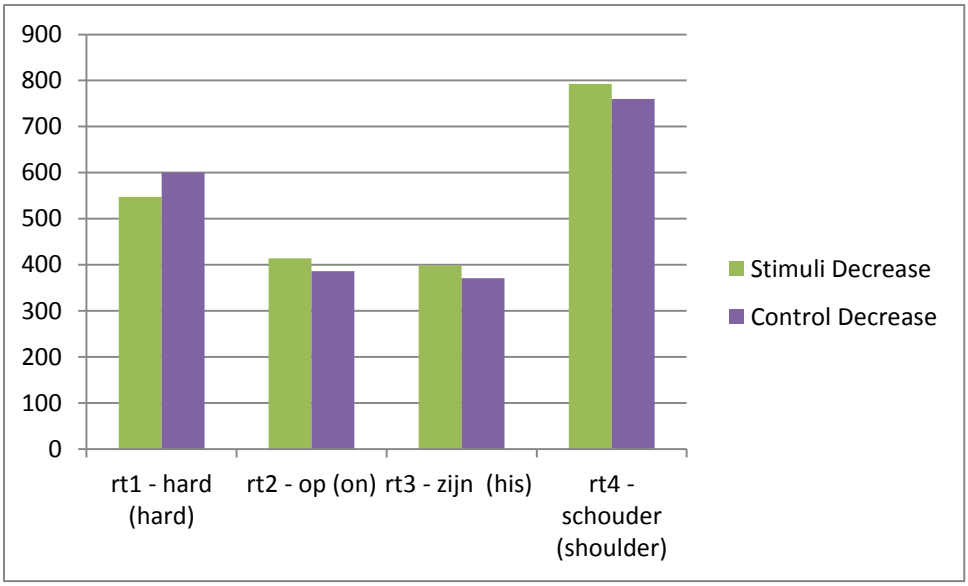


figure 2: mean reading times (decrease condition)

Figure (2) shows the mean reading times of only the decrease condition. These were the sentences such as *Nancy hit Oliver playfully, and simultaneously rather hard, on his shoulder.* Here it can be seen more clearly that the critical adverb, RT1, is read faster in the stimuli sentences than in the control sentences. For every other reading time, the stimuli were read slower than the controls.

Figure (3) shows the mean reading times of only the increase condition. These were the sentences such as *Nancy hit Oliver angrily, and simultaneously rather lightly, on his shoulder.* In all instances, the mean reading time for the stimuli sentences was slower than the mean reading time for the control sentences. Our expectations are thus largely confirmed.

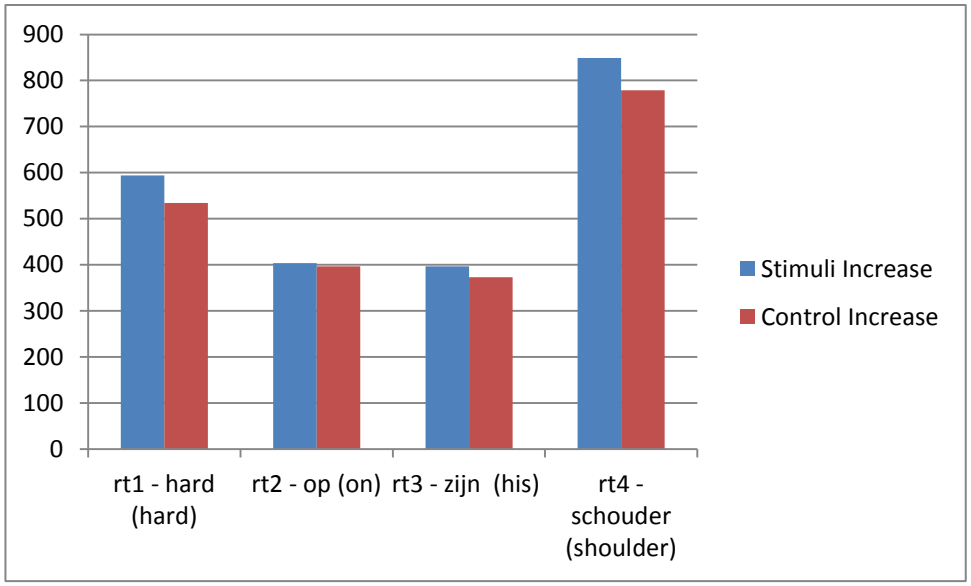


figure 3: mean reading times (increase condition)

To test for the significance of these results, a linear mixed effects analysis on the log-transformed data was done in R. The random effects for this analysis were participants and items. The fixed effects were condition. Results are summarized in Table 1 below.

	RT1 (hard/zacht)	**RT2 (op)**	**RT3 (zijn)**	**RT4 (schouder)**
force increase condition	p = .005** t(56.19) =-2.923 b = 0.072	p = .237 t(57.65) = 1.195 b = 0.023	p = .005** t(56.04) = 2.967 b =0.054	p = .005** t(56.59) = 2.885 b = 0.109
force decrease condition	p = .12 t(56.00) = -1.580 b = -0.046	p = .012* t(57.44) = 2.60 b = 0.059	p = .082 t(55.82) = 1.77 b =0.038	p = .45 t(5.640e+01) = 0.765 b = 3.415e-02

Table 1: p-values, t-values, DOF and standard error (b) for the four reading times

Table (1) gives the p-values, t-values, degrees of freedom and standard estimate (b). Degrees of freedom are noted in between brackets after t; the t-value is noted after the degrees of freedom. Significant differences are marked with one asterisk (*) for $p < 0.05$ or double asterisks (**) for $p < 0.01$. For the increase condition, RT1, RT3 and RT4 are read significantly slower in the stimuli conditions than in the control condition (p=.005); for the decrease condition, only RT2 shows a significant difference (p=.012).

3.2.3 RATING ANALYSIS

An overview of the ratings is given in table (2) below:

ratings/condition	stimuli decrease	control decrease	stimuli increase	control increase
1 (clearly bad)	33	21	124	15
2 (maybe bad)	59	29	195	32
3 (maybe good)	128	86	176	134
4 (clearly good)	208	300	123	425
total	428	436	618	606
Expected ratings	3&4	3&4	2&3	3&4
percentage 3&4	78,5046729	88,53211009	48,38187702	92,24422442

Table 2: total ratings per condition & percentage expected ratings

Table 2 gives the amount of ratings per condition, the total amount of ratings, the expected ratings and the percentage of the total ratings that were on the high end of the scale, i.e. 3 (maybe good) and 4 (clearly good). As can be seen, the control increase sentences received the highest percentage high judgment scores of all ratings (92.2%), followed by the control decrease sentences (88.5%), the decrease stimuli (78.5%) and lastly the increase stimuli (48.4%).

A generalised linear mixed effects model was used on the log-transformed data in R in order to compute the odds of favouring the higher ratings ("clearly good"/"maybe good") over the lower ratings ("clearly bad"/ "maybe bad") between conditions. Ratings 3 "maybe good" and 4 "clearly good" have been pulled together for this analysis, as have been ratings 1 "clearly bad" and 2 "maybe bad"; this allows us to work with binary data without complicated comparisons:

In the increase condition, the odds of choosing the higher ratings (good) over the lower ratings (bad) are 0.913 ($p = 0.673$). These odds are significantly lower than choosing the higher ratings in the decrease condition, which are 20.290 as big ($p < 0.001$); as well as the control

increase condition, which are 13.213 times as big ($p < 0.001$); or the control decrease condition, which are 5.496 times as big, with ($p < 0.001$).

Furthermore, the odds of choosing the higher ratings (good) over the lower ratings (bad) in the decrease condition are 5.017 ($p < 0.001$). These odds are still significantly lower than the odds of choosing the higher ratings over the lower ratings for the control decrease items, which are 2.404 times as big ($p < 0.001$); as well as the odds for the control increase items, which are 3.692 times as big ($p = 0.013$).

Finally, the odds of choosing the higher ratings over the lower ratings in the control increase condition are 18.523 ($p < 0.001$). These odds are significantly higher than the odds for choosing the higher ratings in the control decrease condition, which are 0,651 times as big ($p > 0.05$).

The data in table 2 as well as the odds given above show that the force increase stimuli were judged much worse than all three other conditions. What this means is that participants liked the increase sentences least of them all, the decrease sentences significantly more, but the control sentences more than the either of the stimuli.

3.2.4 DISCUSSION & CONCLUSION

The trend that can be observed provides evidence that there is a reading time delay when an inference is cancelled. This trend was significant in half of the cases. However, force increase stimuli such as *angrily, and at the same time rather lightly* are problematic: the ratings for these items are very diverse, with judgments divided evenly over all four points of the scale. Furthermore, the force increase stimuli received significantly worse ratings than every other condition.

The differences in these ratings are potentially due to prototype effects. The prototypical hitting event between two animate entities involves a high force. This can be seen in the denial-

of-expectation test (cf. chapter 1 & 2) in the sentence *She hit him, but lightly*. The expectation that is denied is one of a high force magnitude, which shows that a prototypical hitting event is carried out with a high amount of force. Our force decrease stimulus *hit friendly, and simultaneously rather hard* thus end up closer to the prototypical hitting event due to *hard* than our force increase stimulus *hit angrily, and simultaneously rather lightly*. This could mean that the inference in our increase stimuli sentences are harder to cancel with verbs of contact by impact, and might even be considered contradictory for some participants.

The same problem does not persist in the decrease stimuli such as *friendly, and at the same time rather hard*. These items ended on the high end of the force magnitude scale as the second adverb in the sentence was always a mono-dimensional adverb that increased force. These items accordingly received ratings that were largely on the high end of the scale (clearly good). Still the force decrease items received significantly worse ratings than both our control conditions.

One explanation for this is the possibility that the stimuli sentences required the cancellation of an inference, or involved a contradiction, whereas control sentences had a 'mutually reinforcing' reading. Instead of reading the sentence in (20) as mildly tautological, both adverbs are reinforcing the force-decrease reading as they are co-directional.

20. Jet sloeg Harry speels, en tegelijk vrij zacht, op zijn schouder.

    Jet hit Harry playfully, and simultaneously rather lightly, on his shoulder.

The self-paced reading experiment above has shown that there is a significant delay in reading time on our multi-dimensional adverbs, and that the cancellation of the inference in multi-attribute adverbs requires additional processing power in the human brain. We take this as evidence that our multi-dimensional adverbs give rise to their inference during on-line processing, and that context does not block them from being triggered (cf. Jaszczolt, 2014).

The multi-dimensional adverbs are therefore a specific class of adverbs that need to be integrated in the current frameworks of semantics. However, our force increasing adverbs

received very poor ratings, which could mean that they contain a contradiction instead of triggering an inference. Additional research is required to examine this effect. In the following section, we will thus leave the force increase items out of consideration and focus on the analysis of the force decrease examples.

4. THEORETICAL IMPLEMENTATION IN NEO-DAVIDSONIAN THEORY

In chapter three we have shown that our multi-dimensional adverbs give rise to an inference that is subsequently cancelled during on-line processing. A theory of adverbs and their different meanings should therefore encompass the possibility of triggering a defeasible inference. This chapter will focus on how to integrate our complex adverbs into an existing Neo-Davidsonian theoretical frame.

4.1 LITERATURE OVERVIEW

In the previous chapters, we have established that there is an inference on adverbs such as *friendly* and *angrily*. Another issue that needs to be discussed is how to accurately represent these inferences in an existing framework. In this section we will first look at some recent literature on manner adverbs to see how they are analysed.

In the introduction, we have seen that there are several dimensions to an event that can be modified by means of a manner adverb, e.g. speed, volume and 'manner'[1] itself (cf. Maienborn & Schäfer, 2011). Furthermore, some adverbs such as *friendly* modify several dimensions. The examples in (5) and (6), repeated here in (21) and (22) show via the denial-of-expectation test that adverbs like *playfully* and *angrily* seem to modify the force dimension as well as the

[1] Keep in mind that the term manner is used in two senses in this thesis: complex manner adverbs characterises a class of adverbs that elucidate how an event took place; 'manner' is a category that is the target of modification for the more complex adverbs that modify a quality dimension which is not as clearly defined as e.g. speed or sound.

'manner' dimension of an event. The expectations that are denied in (21a) and (22a) are that of decreased and increased force, respectively, yet these sentences are still felicitous.

21. a. Nancy hit Oliver playfully, but still rather hard.

    b. # Nancy hit Oliver playfully, but still rather lightly.

    c. # Nancy hit Oliver lightly, but still rather hard.

22. a. Nancy hit Oliver angrily, but still rather lightly.

    b. # Nancy hit Oliver angrily, but still rather hard.

In contrast, the sentences in (21b) and (22b) are infelicitous as they use two adverbs that modify the force dimension in the same direction in contrastive conjunction, i.e. *playfully* and *lightly* or *angrily* and *hard*, while there is no contrast. This shows us that *playfully* and *angrily* do not express a respective decrease/increase in force via their lexical semantics, but only indirectly modify the force dimension by means of an inference.

When compiling a complete formalization of manner adverbs, there is a multitude of facets that should be taken into account. There is extensive literature on the selective pairing of adverbs with verbs (cf. Geuder, 2000, 2006), their context sensitivity, adverbial scope, and the derivation of entailment relations (cf. Schäfer, 2004).

The selective pairing of manner adverbs is shown clearly by Geuder (2000), who shows that a theory of adverbs also needs to take into account that manner adverbs can selectively combine with some verbs, but not with others. Similarly, our multi-attribute adverbs do not give rise to an inference when combined with every type of verb. In (23a), there is a friendly hitting event, which gives rise to the inference that there is little force used. On the other hand, Jet cycling friendly in (23b) has the meaning that it was friendly of Jet to cycle past Stefan's house, not that the manner in which she cycled was friendly. Although cycling requires some sort of effort, combing *friendly* with this motion verb does not give rise to the inference that there was less force exerted in Jet's cycling.

23. a. Jet sloeg Harry vriendelijk op zijn schouder.

    Jet hit Harry friendly on his shoulder.

    b. Jet fietste vriendelijk langs Stefans huis.

    Jet cycled friendly past Stefan's house.

Other facets of manner adverbs is shown in Schäfer (2004), wherein he discusses their context sensitivity, and the derivation of entailment relations. These are exemplified in (24) and (25) respectively.

24. a. Barbara crossed the channel slowly. [in comparison to Hovercrafts]

    b. Barbara crossed the channel quickly. [in comparison to other swimmers]

25. Peter sings loudly → Peter sings.

(Schäfer 2004: 161)

As (24) shows, the manner adverb used to describe the speed of Barbara's swimming varies with the context. In a situation wherein Barbara and several other swimmers are overtaken by a Hovercraft while crossing the channel, both sentences are possible although Barbara's speed does not vary. Sentence (25) illustrates the necessity for a unified theory of (manner) adverbs to also cover the possibility of giving rise to an entailment relation. This entailment relation is shown in sentence (26b), which shows the representation of (26a) in a Davidsonian format (cf. Schäfer 2004).

26. a. Peter sings loudly

    b. ∃e [SING (p, e) & LOUD (e)]

(Schäfer 2004: 162)

27. Peter sings.

However, our multi-dimensional adverbs give rise to an inference, not an entailment (cf. chapter 2). Cancellation of this inference is seen to affect on-line reading times. A complete theory of adverbs should therefore not only cover context sensitivity and entailment relations, but also explain how certain adverbs give rise to an inference in the context of verbs of contact.

The way I intend to do this is via two-level semantics, which has previously been used to capture ambiguities in locative adverbials (cf. Maienborn, 2003; Maienborn & Schäfer, 2011). In the next chapter, I will therefore present the two-level framework that Maienborn (2003) uses on locative prepositions.

4.2 TWO LEVEL APPROACH

Maienborn (2003) starts by stating that one of the merits of the Davidsonian paradigm is its straightforward account of adverbial modification: the paradigm analyses adverbial modifiers as first order predicates that add information about an event. In the typical case of locatives, they specify the location of the referent they modify. See (28):

28. a. Eva signed the contract in the office.

    b. ∃e [SIGN (e) & AGENT (e, Eva) & THEME (e, c) & CONTRACT (c) & LOC (e, IN (o)) & OFFICE (o)]

The semantic form of (28a) is given in (28b), wherein e is a variable that ranges over events, LOC is a relation between individuals (objects or events) and spatial regions, and IN maps objects onto their inner region. However, the same compositional semantics cannot account properly for the following sentence:

29. Eva signed the contract on a separate sheet of paper.

Sentence (29) includes a locative that does not provide information on the location of the whole event, but rather specifies further details about a part of the event, namely the signing. More generally, Maienborn (2003) argues that locative modifiers of the type in (29) express spatial relationships *within* the event described by the verb; they are therefore called event-internal modifiers. Other locatives as in (27) are considered event-external; they provide information about the event argument as a whole (Maienborn 2003: 477).

Maienborn (2003) also shows that the targets of event-internal modifiers are semantically underspecified, and their determination depends largely on world knowledge (2003: 477). Consider the sentences in (30) and (31):

30. Der Koch hat das Hänchen in einer Marihuana-Tunke zubereitet.
    The cook has the chicken in a marijuana sauce prepared.
31. Angela hat sich mit Bardo im Museum verabredet.
    Angela has refl with Bardo in.the museum arranged-to-meet.

(Maienborn, 2003: 479 & 481)

While both sentences are ambiguous to an event-internal or event-external reading, the external reading sentence (30) is rather bizarre. For the external reading of (30), the entire event would take place in a marijuana sauce; whilst preparing the chicken, the cook is standing in a sea of sauce. As this scenario is rather unlikely, world knowledge discards the external reading of the locative modifiers in (30) in favour of the internal reading (Maienborn 2003: 479).

Sentence (31), on the other hand, comes with both readings of the locative modifier as a suitable option: the event-external reading has the making of the appointment taking place in a museum; Angela and Bardo could run into each other while at a museum, and make an appointment to go elsewhere. The event-internal reading specifies the location of the intended appointment; Angela and Bardo have agreed that the location of their meeting will be the museum.

In order to account for the distinction between event-internal and event-external modifiers within the compositional framework, Maienborn (2003) proposes an analysis by means of two-level semantics. She specifies that there are two levels of interpretation: the Semantic Form (SF) and the Conceptual Structure (CS). The semantic form is context independent, the grammatically determined meaning; the conceptual structure is concept-dependent and determined by world knowledge.

Exactly which part of an event the event-internal modifiers relate to depends on context and world knowledge, and is extra-linguistic (Maienborn 2003: 480). The Semantic form of event-internal modifiers does not determine which part of the event is modified, and is therefore underspecified in this respect (2003: 480). This is further substantiated by sentence (32):

32. Maria zog Paul an den Haaren aus dem Zimmer.
    Maria pulled Paul at the hair out of the room.

(Maienborn 2003: 481)

In sentence (32), neither Paul nor Maria are candidates for the entity that is placed at Paul's hair. Our world knowledge tells us that it is most likely Maria's hand, but it is possible for context to determine that it was in fact Maria's teeth, a pair of pincers or something else (Maienborn 2003: 481). Maienborn suggests that this shows that the actual target of an event-internal locative "cannot be determined at the level of SF, where only the grammatically introduced referents are accessible, but must be inferred at the level of CS taking into account context and world knowledge" (2003: 481). In other words, SF cannot identify the target referent for an event-internal modifier; this is left to the conceptual system.

Integrating event-external modifiers is done by means of Maienborn's MOD formula, given in (33). Basically, MOD in (33) merges a modifier with an expression to be modified and yields a conjunction of predicates.

33. MOD: $\lambda Q\, \lambda P\, \lambda x\, [P(x) \,\&\, Q(x)]$

(Maienborn, 2003: 477)

For event-internal modifiers, a different approach is needed. In order to merge a modifier and part of an expression to be modified, Maienborn (2003) introduces an SF-parameter for the located entity in the form of the PART-OF relation. This parameter is introduced as the free variable *v* at the level of SF, and must be instantiated at the level of CS (Maienborn 2003: 487). The formula MOD' is given for this purpose:

34. MOD': λQ λP λx [P(x) & PART-OF (x, v) & Q(v)]

The result of using MOD' to integrate an event-internal modifier is exemplified in (35):

35. Der Bankräuber ist auf dem Fahrrad geflohen.

    The bank robber has on the bicycle escaped.

    a. [$_{PP}$ auf dem Fahrrad]: λx [LOC (x, ON (**b**)) & BIKE (**b**)]

    b. [$_{V}$ geflohen]: λx λe [ESCAPE (e) & THEME (e, x)]

    c. [$_{V}$ [pp auf dem Fahrrad] [$_{V}$ geflohen]]: λx λe [ESCAPE (e) & THEME (e, x) & PART-OF (e, v) & LOC (v, ON (**b**)) & BIKE (**b**)]

(Maienborn 2003: 487)

What the formula in (35) mean is that there is a part *v* of the escaping event *e* that takes place on a the bicycle, but does not specify exactly what part that is in the semantic form. The variable *v* is left unspecified in the semantic form, and can be filled in with contextual knowledge at CS. The knowledge to determine what part of an event the PART-OF relation describes comes from the Conceptual Knowledge Base (CKB).

Maienborn subsequently merges the two very similar templates into MOD*, in the template in (36). This template introduces a free variable *v* and a relational variable R. The condition in (36b) specifies that for event-internal modifiers, based on their syntactic classification (cf. Maienborn 2003), R is a PART-OF relation; in all other cases, i.e. event-external modifiers, R is the identity function.

36. a. MOD*: λQ λP λx [P(x) & R (x, v) & Q(v)]

    b. *Condition on the application of MOD**:

    If Mod* is applief in a structural environment of categorial type X, then R = PART-OF, otherwise (i.e. in an XP-environment) R is the identity function.

(Maienborn 2003: 489)

When looking at the conceptual resolution of the semantic indeterminacy in the compositional semantics of event-internal modifiers, Maienborn (2003) suggests that the SF-parameter for the modified entity must take contextually salient world knowledge into account (2003: 490). In her case, Maienborn argues that event-internal modifiers supply further information about a spatial configuration that is independently established within the conceptual structure (CS) of the event to which they attach. These spatial notions are relevant as conceptual knowledge about event types includes knowledge pertaining to the functional relations that hold between participants. These functional relations are often based on spatial configurations (Maienborn 2003: 490). Maienborn further suggests that event-internal modifiers elaborate on these spatial notions that are part of the verb's CS.

The "conceptual machinery" Maienborn uses for this instantiation is abductive reasoning (Maienborn 2003: 490). This is described as inference to the best explanation; the interpretation of a sentence consists in deriving its most economical explanation that is consistent with what we know. In the case of locatives, Maienborn shows that it takes an underspecified SF and tries to prove it from a conceptual knowledge base (CKB) that "provides axioms, facts and additional contextually legitimated assumptions (Maienborn 2003: 290).

This conceptual knowledge base (henceforth: CKB) is presumed to be common ground between speaker and hearer. Abductive reasoning then leads to a parameter-fixed CS that can "explain" the SF based on the CKB: we try to find a sensible explanation for the underspecified SF by means of backward chaining and factoring (2003: 491). The first of which, backward chaining, is exemplified in (37):

37. $P \rightarrow Q$ *conceptual knowledge*
    Q *underspecified SF*
    ———————
    P *parameter-fixed CS*

(Maienborn 2003: 491)

The way backward chaining works is as follows: conceptual knowledge tells us that P → Q. With the underspecified Q as SF, we try to find a conceptual explanation for Q, and infer that P must be the conceptual background. The "chain" then runs backward from Q to our assumed explanation of P. However, since (X) does not license a valid inference mode, our CKB might license more than one CS explanation for SF, which would then be weighted according to criteria (Maienborn 2003: 491)

Another part of abductive reasoning is factoring: reducing redundancies and leading to a more economical explanation. Factoring "licenses the unification of compatible expressions if the result is consistent with the rest of what is known" (Maienborn 2003: 491). In other words, factoring is simplifying identical variables into one variable as in (38a), simplified by the formula in (38b):

38. a. ∃ … xy … […& P (x) & … & P (y) & …]

    b. ∃ … x … [… & P (x) & …]

(Maienborn 2003: 491)

Parameter fixing then starts with the underspecified SF that needs to be conceptually specified, indicated by SF-parameters. We then try to instantiate these parameters with our conceptual knowledge base by means of backward chaining and factoring. This process is illustrated with the sentence in (35), repeated here as (39a), with the SF in (39b):

39. a. Der Bankräuber ist auf dem Fahrrad geflohen.

    The bank robber has on the bicycle escaped.

    b. SF: ∃e [ESCAPE (e) & THEME (e, **r**) & BANK-ROBBER (**r**) PART-OF (e, v) & LOC(v, ON (**b**)) & BIKE (**b**)]

(Maienborn 2003: 492)

In order to determine the utterance meaning of (39a), some conceptual knowledge is required. To start, Maienborn (2003) proposes that there is shared speaker knowledge of

locomotion, with a distinction between *extrinsic movement* (EXTR-MOVE) and *intrinsic movement* (INTR-MOVE). The first of which, extrinsic movement, involves a vehicle that is used as instrument of locomotion, such as riding and driving; the second, intrinsic movement, is given if part of the object that undergoes movement is used as locomotion, i.e. walking and jumping (Maienborn 2003: 492). This knowledge is captured via the axioms in (40), using the notations *proper part* "⊏" and *mereological difference* "-" (cf. Maienborn 2003):

40. a. ∀exz [MOVE (e) & THEME (e, x) & INSTR (e, z) & VEHICLE (z) & SUPPORT (z, x, τ (e)) → EXTR-MOVE (e)]

    b. ∀exz [MOVE (e) & THEME (e, x) & INSTR (e, z) & z ⊏ x & y=x-z & SUPPORT (z, y, τ (e)) → INTR-MOVE (e)]

    c. ∀ex [EXTR-MOVE] (e) & THEME (e, x) → MOVED-ITEM (e, x)

    d. ∀exyz [INTR-MOVE (e) & THEME (e, x) & INSTR (e, z) & y=x-z → MOVED-ITEM (e, y)

(Maienborn 2003: 492)

The axioms in (40) cover the difference between extrinsic and intrinsic movement. The former involves movement by means of a vehicle that is used as an instrument of locomotion, and said vehicle supports the theme while moving (cf. Maienborn 2003: 492). The latter involves movement in which "part of the object that undergoes movement is used as a means of locomotion" (Maienborn 2003: 492). In the case of extrinsic movement, the item whose movement is dependent on the instrument (MOVED-ITEM) is the theme in (40c); for intrinsic movement, it is the theme minus the bodypart that serves as instrument in (40d) (cf. Maienborn 2003: 492).

Furthermore, the CKB required for Maienborn's (2003) locative prepositions also makes use of the axioms in (41).

41. a. $\forall$e [EXTR-MOVE (e) & ETC_{RIDE}(e) → RIDE (e)]

b. $\forall$e [EXTR-MOVE (e) & ETC_{DRIVE}(e) → DRIVE (e)]

etc.

(Maienborn 2003: 493)

These axioms make use of ETC-predicates, which Maienborn uses as a tool in order to "exploit superset information in the course of abductive reasoning" (Maienborn 2003: 493). When expressing that riding events, for instance, are a subset of extrinsic movement, we would not be able to use this information while backward chaining. However, (41a) converts such an axiom to a bi-directional, which allows us to use this information in either direction: from a superset to a subset, but also from a subset to a superset. This is shown in (42), which follows from the axiom in (41a).

42. $\forall$e [RIDE (e) → EXTR-MOVE (e)]

(Maienborn 2003: 493)

ETC-predicates are then "placeholders for the *differentia specifica* that distinguishes a *species* from its *genus proximum*" (Maienborn, 2003: 493). It might be impossible or improbable to spell them out completely, but they can be assumed by abduction. Maienborn furthermore gives several kinds of axioms required for parameter fixing sentence (39), given in (43) – (46).

43. $\forall$e [INTR-MOVE (e) & ETC_{WALK} (e) → WALK (e)]

44. $\forall$e [X-MOVE (e) & ETC_{ESCAPE} (e) → ESCAPE (e)]

45. a. $\forall$e [EXTR-MOVE (e) → X-MOVE (e)]

b. $\forall$e [INTR-MOVE (e) → X-MOVE (e)]

(Maienborn 2003: 493)

The axiom in (43) gives us an ETC-predicate for intrinsic movement, such as walking or hopping. The axiom in (44) deals with locomotion that can be either extrinsic or intrinsic, such as escaping: one can escape on an island, or on a bicycle. The first of these options is event-

external, the second is event-internal. The axioms in (45) are auxiliary to (44): they specify that an X-movement can be extrinsic or intrinsic.

Finally, in order to use parameter fixing on sentence (39), Maienborn gives us axioms on subkinds of vehicles, as in (46); axioms that relate spatial configurations with functional concepts of containment and support, as in (47); and axioms that to specify event participants as in (48). For further details on these axioms, see Maienborn (2003).

46. $\forall e$ [VEHICLE (x) & ETC_{BIKE} (x) → BIKE (x)]

47. $\forall xyt$ [SUPPORT (x, y, t) & $ETC_{LOC\text{-}ON}$ (y, x) → LOC (yn ON (x))]

48. $\forall ex$ [THEME (e, x) → PART-OF (e, x)]

(Maienborn 2003: 494)

In total, the axioms supply a suitable background for using abductive interpretation on sentence (39). Applying backward chaining and factoring to our initial SF in (39b) gives us the possibility of identifying *v* as the discourse referent *der Bankräuber*. A slightly adapted figure of how this works is taken from Maienborn (2003:495) and given in figure (4). In this figure, the numbering that Maienborn uses to label her axioms has been replaced with the numbers 1-6 in order to accurately relate them to the axioms as represented in this thesis.

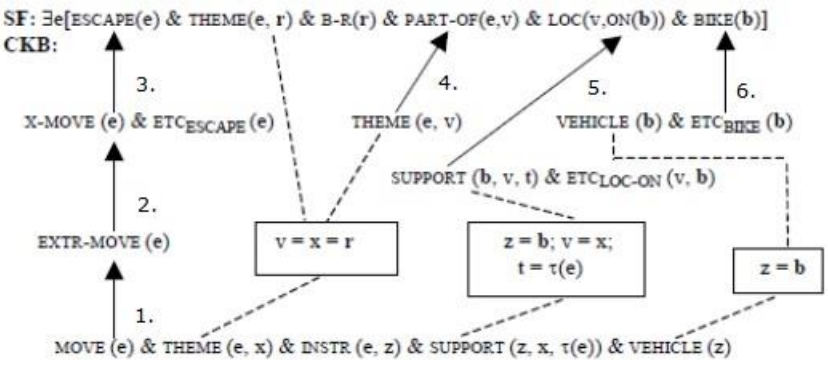


figure 4: backward chaining and factoring at work

(Maienborn 2003: 495)

Figure (4) shows how the relevant axioms, represented by the full arrows. Factoring is indicated by means of broken lines that link equations to the relevant literals. Arrow (3) uses the axiom in (44) that says escaping is a type of X-MOVE, which can be intrinsic or extrinsic. The second arrow, (2), depicts the axiom in (45a), which says that this type of X-MOVE is extrinsic. Arrow (1) depicts the axiom in (40), stating that the extrinsic movement is a type of movement. The other arrows (4), (5) and (6) indicate the axioms in (48), (47) and (46), respectively. For further details on these axioms, see Maienborn (2003: 494).

Once backward chaining and factoring has occurred based on the axioms in our CKB, the parameter-fixed CS for the sentence is derived. Abductive reasoning then proceeds to prove the underspecified SF from the CS, leading to the evidence that CS is in fact a possible specification of the underspecified SF with respect to our CKB.

4.3 A FIRST APPROACH

As manner adverbs are standardly treated as predicates of events, I present an example analysis of the adverb *friendly* modifying the verb *hit* in Neo-Davidsonian terms, in the spirit of Maienborn's (2003) two-level semantics. The assumption is that this analysis can be extended to similar adverbs, i.e. force decreasing multi-dimensional manner adverbs, and other verbs, i.e. verbs of contact by impact.

As discussed in chapter 1 above, the class of adverbs that has previously been called "manner adverbs" can be further divided into "true" manner adverbs, compatible with the *in an X manner* paraphrase, and adverbs modifying some other dimension of the event, e.g. the sound-volume or force. Given the evidence from the *in an X manner*-paraphrase, we assume that events that can be modified with (a subset of) other dimensions, such as sound-volume or force, come equipped with a corresponding 'manner', sound-volume or force component in their lexical entries. For example, the adverb *hard* can modify the verb *hit* because hitting has a force

component, and *loudly* can modify the verb *speak* because *speak* has a sound-volume component. In these examples, switching the modifiers is not easily done: *to hit loudly* or *to speak hard* sounds odd. In accordance with McConnel-Ginet (1982), we assume that verbs like *run* have a latent argument slot for speed, which may be activated and filled in by an adverb such as *quickly* (cf. McConnel-Ginet, 1982, 169-170); and that verbs like *hit* have a latent argument slot for force. This leads to our proposal for the lexical entry for *hit* as the following, cf. (49):

49. hit = λy.λx.λe.∃f [hit(e) & AGENT(e, x) & PATIENT(e, y) & ∃m [MANNER(e, m)] & ∃f [FORCE(e, f) & magnitude(f, v)]]

Contra Schäfer (2015), who introduces manner via a manner template instead of in the lexical entry of the verb, (49) specifies the lexical entry for hit with both a manner and a force component. Note that while we have specified the event for force of a certain magnitude, the exact value of that magnitude remains underspecified. This underspecification is depicted in the unbound variable *v* with a parameter that needs to be filled in, but is dependent on context and world-knowledge; this is based on the two-level idea from Maienborn (2003).

The variable *v* needs to be filled out by our Conceptual Knowledge Base (CKB) (cf. Maienborn 2003). Although this is not represented in the formula in (49), we have reason to believe that the default assumption for a hitting between two animate entities is that there is a high force component involved. The denial-of-expectation test (cf. Lakoff, 1971) in sentence (50a) versus the infelicitous (50b) shows us that the expectation for the verb *hit* is that there is a high force involved. However, the availability of *lightly* as a modifier of *hit* shows us that the force magnitude can also receive a low value, but that this is not the prototypical case.

50. a. She hit him, but lightly.
    b. # She hit him, but hard.

The lexical semantics of *lightly* and *hard* are notoriously difficult to define. In this thesis, I do not want to make any claim here as to how the polysemy of these adverbs is best treated. Instead,

simplified lexical entries are given that contain meaning components which are part of the lexical semantics of these words. These simplified lexical entries can be further extended in order to accommodate more meaning aspects. In accordance with a.o. Piñón (2009) and Maienborn (2003), we treat *lightly* and *hard* as predicates over events, with the force component and the magnitude specified in their lexical entry, cf. (51).

51. a. lightly = λe.∃f [FORCE(e, f) & magnitude(f, low)]

    b. hard = λe.∃f [FORCE(e, f) & magnitude(f, high)]

Finally, the lexical entry for manner adverbs such as *friendly* needs to specify the manner component, cf. (36). Following Schäfer (2013), the manner modifier in (52) predicates over the manner variable, which is related to the event variable by a manner function (2013: 191). As with *hard* and *lightly*, a simplified version is given that leaves open the possibility of extending the lexical entry with other meaning aspects when required.

52. friendly = λe. ∃m [MANNER(e, m) & friendly (m)]

With the lexical entries spelled out, we now turn to the compositional semantics, to generate the corresponding SF for our example sentences. Given our lexical entries for *hit*, *hard*, and *lightly*, the sentence in (53a) receives the sentence representation in (53b):

53. a. Nancy hit Oliver lightly.

    b. ∃e [hit (e) & AGENT (e, nancy) & PATIENT (e, oliver)& ∃m [MANNER (e, m) & friendly (m) & ∃f [FORCE (e, f) & magnitude (f, low)]]

The verb *hit* in (53) is equipped with a force component, which has a magnitude that is left unspecified. The specification of the force magnitude is contributed by the adverb *low*, resulting in a hitting event that is done with low force.

For our multi-dimensional adverbs, on the other hand, the analysis is less straightforward, shown in sentence (54a) with the logical form in (54b):

54. a. Nancy hit Oliver friendly.

b. ∃e [hit(e) & AGENT(e, Nancy) & PATIENT(e, Oliver) & ∃m [MANNER(e, m) & friendly(m)] & ∃f [FORCE(e, f) & magnitude(f, v)]]

The logical form in (54b) is slightly more complex than the one in (53b). In (54b), the magnitude of the force remains underspecified, as can be seen by the unbound variable *v* in the logical form. This means that the magnitude of the force of the hitting event is not specified at SF in this case, and will be specified at the level of CS.

At the level of CS, the unbound variable *v* in (54b) needs to be replaced with the value *low* in order to derive the correct (defeasible) inference. Following Maienborn (2003), I use abductive interpretation as a formal means of parameter fixing, which is based on meaning axioms coded in the Conceptual Knowledge Base (CKB), These axioms contain information about e.g. the event participants, and in our case social conventions.

In order to account for (54b), a meaning axiom is required such as in (55), stating that for all events and manners, if the event has a force and the manner is friendly, the magnitude of the force will be low. This axiom is based on the movement axioms in (40) in section 3.2, taken from Maienborn (2001), but uses the angled bracket > to represent a defeasible implication, as in Asher and Morreau (1991).

55. ∀e∀m [[manner(e, m) & friendly(m) & ∃f [force(e, f)]] > magnitude (f, low)]

The axiom in (55) and other similar axioms accounting for the force decreasing effects of multi-dimensional adverbs such as *friendly* and *playfully* are based on a general underlying mechanism. We assume there is a translation from a phrase such as *hit friendly* to the conceptual level of underlying meanings of social conventions and interactions. From this level, concrete

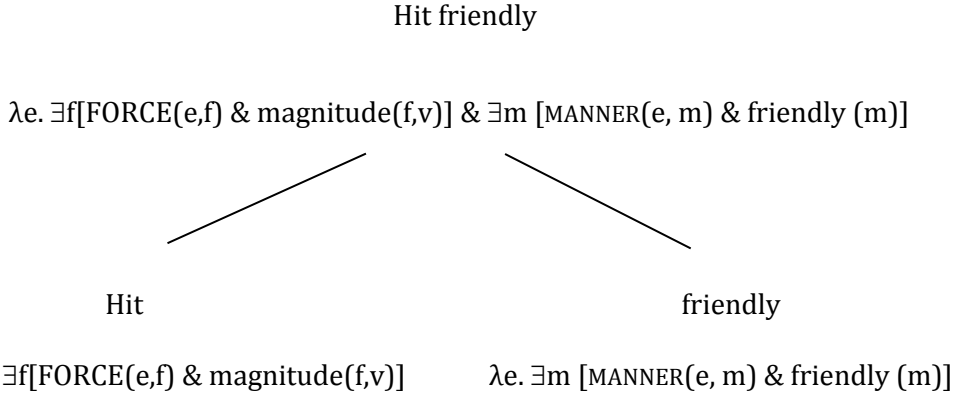
manifestations of physical behaviour can be derived, such as typically hitting with low force if one is hitting in a friendly manner.

At first glance, our first attempt at explaining our inference via Maienborn's (2003) two-level approach explains the phenomenon beautifully. Unfortunately, we run into a problem upon further inspection. While this framework can explain the inference on adverbs such as *friendly*, it does not deal equally well with mono-dimensional adverbs such as *lightly*. To illustrate this, we use a simplified lexical entry for *hit* as in (56), omitting information about punctual contact. This lexical entry still states that a hitting event involves force, and the force is underspecified in the variable *v*.

56. hit = λe… ∃f … ∃m [MANNER(e, m) & ∃f [FORCE(e, f) & magnitude(f, v)]]

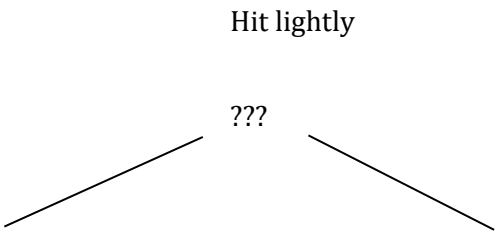
At the level of SF, all the chunks that can be combined, are combined. Therefore, given the lexical entries above, the composition of the verb *hit* and the manner adverb *friendly* will look as follows:

57.



In the derivational process, *v* is left underspecified in *hit friendly*, which is exactly what we want since it will be specified at the level of CS. However, the composition of the verb *hit* and the manner adverb *lightly* receives a similar analysis:

58.



The combination of *hit* and *lightly* is intentionally left into question here, as both *hit* and *lightly* have a magnitude element. For *lightly* this is specified; the assumption was that this would specify the unbound variable *v* in the magnitude of *hit*. However, if we stay faithful to Maienborn, in order for this to happen *v* needs to pass through the second level, the CS, in order to be interpreted. In other words, the interpretation of our unbound variable does not happen until the level of Conceptual Structure, as opposed to the level of Semantic Form.

4.4 MAIENBORN EXTENDED

A solution for our problem with Maienborn's (2003) two-level approach lies in postulating that her unbound variable *v* can be filled in at any time during the derivation. In other words: an option to fix the compositional problem is to regard *v* as a placeholder that can be kicked out in any place in the composition. This is in the spirit of Maienborn, but not her exact formulation.

This still leaves us with the unbound variable *v* inserted at the level of SF, functioning as a signal for CS that further specification by context or world knowledge is required in order to fully interpret this part of the sentence. If we assume that the unbound variable *v* can still be replaced with an inferred force magnitude at the level of CS, we can take a closer look at the components of our CKB. For this, we will require a further description of axioms for verbs of

contact by impact, axioms for the event participants and axioms for social situations described by *friendly* and other force decreasing adverbs.

Verbs of contact by impact are sub-classified by Levin (1993) into *hit* verbs, *swat* verbs, and *spank* verbs. All three are described as moving one entity in order to bring it into contact with another entity, but do not necessarily entail that this contact has any effect on the second entity (Levin, 1993: 151). For syntactic rules related to the alteration between the three types of verbs of contact by impact, see Levin (1993).

In order to extend our example analysis of hit-type verbs to other contact-by-impact verbs, we make use of an ETC-predicate. These predicates, as explained in section 4.2, can exploit the superset information in the course of abductive reasoning by converting axioms into biconditionals. In other words, the axiom in (59a) makes use of an ETC predicate so that we are able to express that hitting events are a subset of verbs of contact by impact, as in (59b), while making use of backward chaining. With the ETC-predicate, we are able to use this axiom in both directions. Our analysis of the *hit* verb then extends to the larger category of contact-by-impact verbs. For this, we need the axiom in (59a) (cf. Maienborn 2003: 493).

59. a. $\forall$e [CONTACT-IMPACT (e) & ETC_{HIT}(e) → HIT (e)]

    b. $\forall$e [HIT (e) → CONTACT-IMPACT (e)]

Furthermore, I propose another axiom in order to specify that friendly manners are a sub-type of manners, and more specifically a sub-type of what I will call social manners (SOC-MAN). In order to accurately describe axioms for social manners, we make us of an ETC-predicate over manners, similar to those in sentence (41) (cf. chapter 4.2). This allows us to use subset information, namely friendly, as a basis for backward chaining. The ETC-predicate for social manners, or SOC-MAN, is given in (61).

60. ∀e∀m [[MANNER(e, m) & FRIENDLY(m) & ∃F [FORCE(e, f)]] > MAGNITUDE (f, low)]

61. a. ∀m [SOC-MAN (m) & $ETC_{FRIENDLY}$ (m) → FRIENDLY (m)]

b. ∀m [FRIENDLY (m) → SOC-MAN (m)]

We then slightly alter our axiom in (55) to form the one in (62), which states that social manners combined with events that involve force lead to the defeasible inference of low force magnitude. This allows us to say that the inference that there is low force magnitude does not stem from the adverb *friendly* but is in fact a case of social manners such as *friendly*, *playfully* and other adverb belonging to that category. The angled bracket > represents our defeasible inference (cf. Asher & Morreau, 1991).

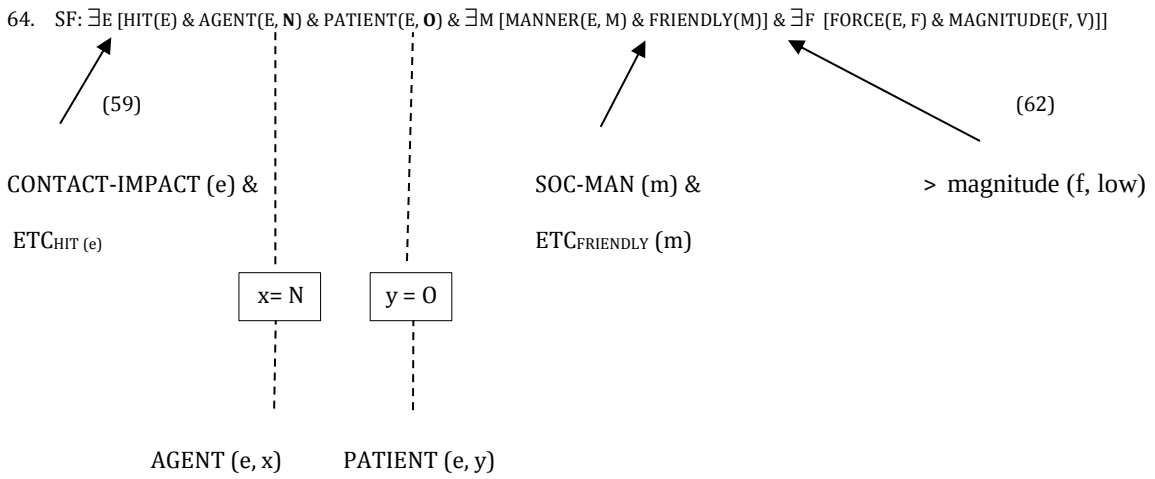
62. ∀e∀m [[manner(e, m) & SOC-MAN (m) & ∃f [force(e, f)]] > magnitude (f, low)]

The axioms in (59) – (62) provide a background for the abductive interpretation of the sentence in (63a). The initial SF is given in (63b). Proper names are abbreviated to their initials so that **n** stands for Nany and **o** stands for Oliver.

63. a. Nancy hit Oliver friendly.

b. ∃e [HIT(e) & AGENT(e, **n**) & PATIENT(e, **o**) & ∃m [MANNER(e, m) & FRIENDLY(m)] & ∃f [FORCE(e, f) & MAGNITUDE(f, v)]]

Through the processes of backward chaining and factoring, the initial SF in (63b) yields a possible CS that gives a low magnitude for the SF-parameter *v*. This is illustrated in (64) below, with the relevant axioms noted between brackets next to the arrows. Factoring is indicated by equations in squares that are linked to their literals by broken lines. The inference on friendly is indicated by the angled bracket ">".



Following the backward chaining and factoring in (64), we get a parameter-fixed CS shown in (65). This CS is a plausible utterance meaning for the sentence in (63a).

65. ∃E [CONTACT-IMPACT (e) & ETC_{HIT}(e) & AGENT(e, n) & NANCY (n) & PATIENT(e, **o**) & OLIVER (o)& ∃m [MANNER(e, m) & FRIENDLY(m) & $ETC_{FRIENDLY}$ (m)] & ∃f [FORCE(e, f) & MAGNITUDE(f, low)]]

Keep in mind, however, that it is only one of the possible meanings. The variable *v* is still only filled out by means of an inference, indicating the defeasibility of the effect that *friendly* has on the force magnitude. When sentence (63a) is combined with an adverb in an opposite force direction in a sentence such as *Nancy hit Oliver friendly and simultaneously rather hard*, the inference is cancelled and the force magnitude will be specified via the mono-dimensional adverb *hard*.

To conclude, this chapter presents an example analysis of the multi-dimensional adverb *friendly* in the context of *hit*. By means of Maienborn's (2003) two-level semantics we have shown how a possible Conceptual Knowledge Base (CKB) can be used as a semantic background for computing inferences on the force dimension.

5. CONCLUSION

The purpose of this thesis was to find empirical evidence for the subdivision of manner adverbs into mono-dimensional adverbs such as *lightly*, directly modifying a quality dimension of the verb; and adverbs such as *friendly*, that target a more complex dimension; and to incorporate these multidimensional adverbs in current Neo-Davidsonian theory. In chapter two we have seen that multi-dimensional manner adverbs trigger a defeasible inference on the force dimension, decreasing the magnitude, when used to modify verbs of contact by impact.

Evidence for this type of inference is presented in chapter three in the form of questionnaires and a self-paced reading study. Results of the self-paced reading study showed a significant reading time delay after cancellation of the inference, which means that the process of cancellation requires additional processing power in the human brain. I take this as evidence that our multi-dimensional adverbs give rise to their inference during on-line processing, and that context does not block them from being triggered. Furthermore, I take this to mean that the presumptive meanings discussed in chapter 1 do arise, and are subsequently cancelled (cf. Jaszczolt, 2014).

However, sentences which contained adverbs such as *angrily*, which we describe as multi-dimensional in chapter 2, show questionable judgment scores. This poor judgment scores are potentially due to prototype effects. As illustrated by the denial-of-expectation in the sentence *She hit him, but lightly*, a prototypical hitting event between two animate entities involves a high force (cf. chapter 1 and 2): the expectation that is denied is one of a high force magnitude. Our force decrease stimuli thus end up closer to a prototypical hitting event, as they end on a force-increasing adverb. This could mean that our force-increasing stimuli have a slightly contradictory nature, and are therefore rated lower. However, in order to ascertain this, further research is required to define why force-increasing multi-dimensional adverbs are more problematic than force-decreasing multi-dimensional adverbs.

An example analysis of the multi-dimensional adverb *friendly* within the current framework of Neo-Davidsonian Event Semantics was presented in chapter 4. This analysis made use of a two-level approach as used by Maienborn (2003), with the levels of Semantic Form and Conceptual Structure. The magnitude of the force is left unspecified at the level of compositional meaning (SF). Subsequently, the magnitude is computed at the conceptual level by means of meaning axioms about social conventions, which are part of our Conceptual Knowledge Base (CKB). In this example analysis, I make use of ETC-predicates in order to say that the inference does not stem from the adverb friendly alone, but is possible for a larger group of adverbs that describe social situations. Similarly, this inference can arise not only with *hit* verbs but also with other verbs of contact by impact. This is also formulated in an axiom via an ETC-predicate, based on previous work by Maienborn (2003).

These ETC-predicates license backward chaining from multi-dimensional adverbs such as *friendly* to the superset information of what I have called social manners. These social manners have in common that they trigger a defeasible inference of a low force magnitude when combined with verbs of contact by impact, such as *hit*. However, what is missing is a discussion of how these social manners fit within the general framework of manners, and what kind of manners multi-dimensional force-increasing adverbs such as *angrily* belong to. Future research can attempt to provide an answer to these questions.

In conclusion, although the case study presented here is only a small study in the domain of forces, it shows the importance of conceptual knowledge for natural language meaning and utterance interpretation.

6. BIBLIOGRAPHY

Asher, Nicholas, and Michael Morreau. 1991. "Commonsense Entailment: A Modal, Nonmonotonic Theory of Reasoning." *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence* ed. by John Mylopoulos and Ray Reiter. San Mateo, Calif.: Morgan Kaufmann.

Carston, R. (2004). Stephen C. Levinson, Presumptive meanings: the theory of generalized conversational implicature. Cambridge, MA: MIT Press, 2000. Pp. xxiii+ 480. *Journal of linguistics*, *40*(1), 181-186.

Geuder, W. (2006). Manner modification of states. In: Ch. Ebert & C. Endriss (eds.). Proceedings of Sinn und Bedeutung (= SuB) 10. Berlin: ZAS, 111–124

Goldschmidt, A. (2013). On lightly" and playfully" and other such manners. Generals Paper - Universiteit Utrecht, UiL OTS.

Haugh, M. (2013). Implicature, inference and cancellability. In *Perspectives on Pragmatics and Philosophy* (pp. 133-151). Springer International Publishing.

Hofland, S.B. (2011). *Adverbs in Conceptual Semantics.* University of Oxford.

Horn, L. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. *Meaning, form, and use in context: Linguistic applications*, 11-42.

Horn, Laurence R. (2004). "Implicature". *Handbook of Pragmatics* ed. by Laurence R. Horn & Gregory Ward, 3-28. Oxford: Blackwell.

Jaszczolt, Katarzyna M. (2014). Defaults in Semantics and Pragmatics. In: The Stanford Encyclopedia of Philosophy (Winter 2014 Edition), Edward N. Zalta (ed.)

Koons, R. (2014). Defeasible Reasoning. In: The Stanford Encyclopedia of Philosophy (Spring 2014 Edition), Edward N. Zalta (ed.),

Lakoff, R. (1971). "If 's, and's and but's about Conjunction". *Studies in Linguistic Semantics* ed. by Charles J. Fillmore & Donald T. Langedoen, 114-149. New York: Holt, Rinehart and Winston, Inc.

Levin, B. (1993). English verb classes and alternations - a preliminary investigation. Chicago and London: The University of Chicago Press, 1 edition.

Levinson, S. C. (1995). Three levels of meaning. In *Grammar and meaning: Essays in honour of Sir John Lyons* (pp. 90-115). Cambridge University Press.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT press.

Maienborn, C. (2003). Event-internal modifiers: Semantic underspecification and conceptual interpretation. In E. Lang, C. Maienborn & C. Fabricius-Hansen (eds): Modifying Adjuncts (Interface Explorations 4), 475-409.

Maienborn, C., and Martin Schäfer. (2011). Adverbs and adverbials. In Semantics. an international handbook of natural language meaning, ed. Claudia Maienborn, Klaus von Heusinger, and Paul Portner, volume 1 of HSK Handbook Series, chapter 54. Berlin, New York: Mouton de Gruyter.

McConnell-Ginet, S. (1982). "Adverbs and Logical Form: A Linguistically Realistic Theory". *Language* 58.144-184.

Piñon, C. (2009). Agent-oriented adverbs as manner adverbs . Handout, *Ereignissemantik-Workshop*, Humboldt-Universität, Berlin, 11–12 Dec. 2009.

Schäfer, M. (2004). Manner adverbs and events: resolving the underspecifcation. In: Asatiani, Balogh et. al. (eds.). Proceedings of the Fifth Tbilisi Symposium on Language, Logic and Computation". Institute for Language, Logic and Computation, University of Amsterdam: 161-166.

Schäfer, M. (2013). *Positions and interpretations: German adverbial adjectives at the syntax-semantics interface* (Vol. 245). Walter de Gruyter.

Schäfer, M. (2015). What do manner adverbs modify? Handout. Leiden Utrecht Semantics Happenings, Universiteit Utrecht, 26 Nov. 2015.